International Journal of Computer Science Engineering and Information Technology Research (IJCSEITR)
ISSN(P): 2249-6831; ISSN(E): 2249-7943
Vol. 4, Issue 2, Apr 2014, 237-244

DETERMINING AND BLOCKING OF SYBIL USERS ON ONLINE SOCIAL NETWORK

MAHENDRA EKNATH PAWAR[1] & B. W. BALKHANDE[2]

[1]Department of Computer Engineering, Bharati Vidyapeeth College of Engineering, Navi Mumbai, Maharashtra, India

[2]Professor, Department of Computer Engineering, Bharati Vidyapeeth College of Engineering, Navi Mumbai, Maharashtra, India

ABSTRACT

Sybil attacks are one of the well-known and powerful attacks against OSNs. The malicious attacker generates a Sybil group, and pretends to be multiple, distinct users (called Sybil users). Sybil attacks have several harms in OSNs. Dangerous is that multiple Sybil users collude together and form a Sybil group. In this project, we present the first attempt to identify and validate Sybil groups in online social network. We build Sybil group detector based on multiple attributes. We present the first attempt to identify and validate Sybil groups in the real system. First of all, we build the Sybil group detector based on multiple attributes, Including activity, popularity, social degree, friend relationship and IP address. Our Sybil group detection and validation mechanisms have important implications for system design to defend against Sybil attacks in OSNs.

KEYWORDS: Detection, Validation, Online Social Networks

INTRODUCTION

Kinship is a central area of study for social a social network is a social structure made up of a set of social actors (such as individuals or organizations) and a set of the dyadic ties between these actors. The social network perspective provides a set of methods for analyzing the structure of whole social entities as well as a variety of theories explaining the patterns observed in these structures. The study of these structures uses social network analysis to identify local and global patterns, locate influential entities, and examine network dynamics.

A social networking site is a platform to build social networks or social relations among people who, for example, share interests, activities, backgrounds, or real-life connections. A social network service consists of a representation of each user (often a profile), his/her social links, and a variety of additional services. Most social network services are web-based and provide means for users to interact over the Internet, such as e-mail and instant messaging. Social networking sites allow users to share ideas, pictures, posts, activities, events, and interest with people in their network.

In recent years, online social networks (OSNs) become huge and they are still growing throughout the world. Unfortunately, the relative openness and the tremendous growth of OSNs attract the interest of malicious parties.

Sybil attacks are one of the well-known and powerful attacks against OSNs. The malicious attacker generates a Sybil group, and pretends to be multiple, distinct users (called Sybil users). Sybil attacks have several harms in OSNs. First of all, multiple Sybil users are utilized to unfairly increase influence and power of target users. For example, commercial services use many Sybil users and promote clients' content to the top position in Youtube1. In Twitter, some

political campaigns disguise themselves as spontaneous grassroots behavior that is actually carried out by a single organization. In Facebook, gamers control Sybil users to achieve higher status in social games. Secondly, spammers target OSNs as media to propagate spam. In Facebook, compromised accounts send malicious wall posts with embedded URLs. Malicious users rely on unsolicited mentions or embedding hash tags to send spam content in Twitter. Sybil attacks become increasingly dangerous as more people use OSNs as primary interfaces to the Internet.

Successfully defending against Sybil attacks is important to ensure fairness and credibility in the system, to reduce user burden of dealing with spam, and to positively impact the overall value of OSNs going forward. Previous works utilize friend relationships to detect Sybil users, including Sybil-Guard, Sybil Limit, Sybil Infer, Sum Up and the Sybil detector.

Initial studies focus on detecting Sybil users. However, Sybil users alone do not harm the system. What is really dangerous is that multiple Sybil users collude together and form a Sybil group. The attacker controls the Sybil group to attack the system seriously. However, few studies have analyzed the relationship between Sybil users and detected Sybil groups in OSNs. Identifying Sybil groups can be used to detect attackers, who create and control Sybil users. A further step can be taken to study behaviors of attackers, and design new mechanism to prevent attacks.

We design automatic validation mechanisms of Sybil groups, by analyzing action time similarity of users in a group. We apply the validation methods to Sybil groups. We observe that users in Sybil groups show extremely high similarity of action time than that of users in normal groups. It indicates that Sybil users are simultaneously controlled by the same attacker. In summary, we present the first attempt to identify and validate Sybil groups in online social networks. We utilize multiple attributes to detect Sybil users and identify Sybil groups in the real system. Our results are confirmed by automatic validation mechanisms, instead of simulation experiments or manual inspections. Our Sybil group detection and validation mechanisms have important implications for system design to defend against Sybil attacks in OSNs.

RELATED WORK

Due to the explosive growth and popularity of social networking communities, a great deal of research has been done to study various aspects of these communities. Specifically, these studies have focused on usage patterns [6, 10], information revelation patterns [6, 14], and social implications [7, 8] of the most popular communities. Work has also been done to characterize the growth of these communities [16] and to predict new friendships [17] and group formations [3]. Recently, researchers have also begun investigating the darker side of these communities. For example, numerous studies have explored the privacy threats associated with public information revelation in the communities [4, 11]. Aside from privacy risks, researchers have also identified attacks that are directed at these communities (e.g., social spam) [13].

In our previous work [3], we showed that social networking communities are susceptible to two broad classes of attacks: traditional attacks that have been adapted to these communities (e.g., malware propagation) and new attacks that have emerged from within the communities (e.g., deceptive spam profiles). Unfortunately, very little work has been done to address the emerging security threats in social networking communities. However, the research community desperately needs real-world examples and characterizations of malicious activity to inspire new solutions. Thus, to help address this problem, we present a novel technique for collecting deceptive spam profiles in social networking communities that relies

on social honey pot profiles. Additionally, we provide the first characterization of deceptive spam profiles in an effort to stimulate research progress.

The Facebook Social Network

Facebook is currently the largest social network on the Internet. On their website, the Facebook administrators claim to have more than 400 million active users all over the world, with over 2 billion media items (videos and pictures) shared every week [3].

Usually, user profiles are not public, and the right to view a user's page is granted only after having established a relationship of trust (paraphrasing the Facebook terminology, becoming friends) with the user. When a user A wants to become friend with another user B, the platform first sends a request to B, who has to acknowledge that she knows A. When B confirms the request, a friendship connection with A is established. However, the users' perception of Facebook friendship is different from their perception of a relationship in real life. Most of the time, Facebook users accept friendship requests from persons they barely know, while in real life, the person asking to be friend would undergo more scrutiny.

In the past, most Facebook users were grouped in networks, where people coming from a certain country, town, or school could find their neighbors or peers. The default privacy setting for Facebook was to allow all people in the same network to view each other's profiles. Thus, a malicious user could join a large network to crawl data from the users on that network. This data allows an adversary to carry out targeted attacks. For example, a spammer could run a campaign that targets only those users whose profiles have certain characteristics (e.g., gender, age, interests) and who, therefore, might be more responsive to that campaign. For this reason, Facebook deprecated geographic networks in October 2009. School and company networks are still available, but their security is better, since to join one of these networks, a user has to provide a valid e-mail address from that institution (e.g., a university e-mail address).

The MySpace Social Network

MySpace was the first social network to gain significant popularity among Internet users. The basic idea of this network is to provide each user with a web page, which the user can then personalize with information about herself and her interests. Even though MySpace has also the concept of "friendship," like Facebook, MySpace pages are public by default. Therefore, it is easier for a malicious user to obtain sensitive information about a user on MySpace than on Facebook. Users might be profiled by gender, age, or nationality, and an aimed spam campaign could target a specific group of users to enhance its effectiveness. MySpace used to be the largest social network on the Internet. Although it is steadily losing users, who are mainly moving to Facebook [2], it remains the third most visited site of its kind on the Internet.

The Twitter Social Network

Twitter is a much simpler social network than Facebook and MySpace. It is designed as a micro blogging platform, where users send short text messages (i.e., tweets) that appear on their friends' pages. Unlike Facebook and MySpace, no personal information is shown on Twitter pages by default. Users are identified only by a username and, optionally, by a real name. To profile a user, it is possible to analyze the tweets she sends, and the feed to which she is subscribed. However, this is significantly more difficult than on the other social networks. A Twitter user can start "following" another user. As a consequence, she receives the user's tweets on her own page. The user who is "followed"

can, if she wants, follow the other one back. Tweets can be grouped by hash tags, which are popular words, beginning with a "#" character. This allows users to efficiently search who is posting topics of interest at a certain time. When a user likes someone's tweet, he can decide to rewet it. As a result, that message is shown to all her followers. By default, profiles on Twitter are public, but a user can decide to protect her profile. By doing that, anyone wanting to follow the user needs her permission. According to the same statistics, Twitter is the social network that has the fastest growing rate on the Internet.

PROBLEM STATEMENT

Initial studies characterize the spam problem in online social networks. Previous approaches use blacklists to identify malicious URLs in Facebook [3] and Twitter [4]. Ratkiewicz et al. detect the early stages of viral spreading of political misinformation in Twitter [2]. These works mainly detect spam messages, and analyze malicious behaviors. In contrast, we focus on identifying and validating sybil groups, rather than spam content. Various techniques are applied to study sybil users or spammers in OSNs. First of all, several Sybil defense schemes [12] are based on the assumption that sybil users can hardly make friends with normal users [15], [16]. Secondly, honeypots are deployed to trap spammers who attempt to make friends with them in Twitter [18] and Myspace [19]. Thirdly, researchers manually identify spam tweets in Twitter [20], phantom profiles in Facebook [1] and spammers in Youtube [21]. Finally, Thomas et al. identify accounts suspended by Twitter for disruptive activities [7]; Yang et al. analyze friend requests to detect sybil users [13]; Yardi et al. examine spam around the Twitter meme to detect spammers [22]. Our works are different from these studies in several fields: first of all, previous works identify sybil users or spammers, without detecting malicious groups. We analyze the relationship between sybil users, and further identify sybil groups. Secondly, some initial studies mainly use simulation experiments or manual inspections to verify their methods; other works verify sybil users only after spam content has been posted or abnormal users have been forbidden. We apply our sybil group detector to the large-scale datasets in Online social network site, and design automatic mechanisms to validate our detector in the real system. Malicious behaviors are not necessary in our validation methods, and potential attacks can be prevented beforehand.

PROPSED SYSTEM

Determining Sybil Users

Relationship between Popularity and Social Degree

We begin by giving the definition of some properties in social network sites. People establish bidirectional social relationships with friends in OSNs. *Social degree* is defined as the number of friends. Standard user is limited to a maximum of 1,000 friends in Online social network site. Users may pay a subscription fee to increase this limit to 2,000. *Popularity* is defined as the number of visits a user's profile receives [14]. The popularity reflects how attractive the user's profile is. The popularity always increases as the social degree grows. The popularity sometimes grows as the social degree remains unchanged. The popularity is the number of profile browsing by all visitors. A significant of visitors are strangers [14]. Strangers increase the popularity but have no contribution to the social degree.

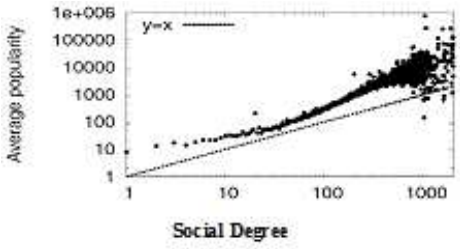

Figure 1: Social Degree versus Average Popularity

Generally speaking, the user's popularity is at least equal to he social degree. We measure the relationship between the popularity and the social degree in the real system. We obtain 1,000,000 random users' popularity and social degree. We compute the average popularity of users who have the same social degree. Then we plot the social degree versus average popularity in Figure 1. As the growth of social degree, the average popularity also rises up. Most of points are above the line $y = x$, and the average popularity is higher than the social degree. Only 8 points are below the line $y = x$. We manually check 8 points, and find them abnormal.

Tracking Using Login Time

The last login time describes the time when the user logins into the OSN for the last time. For a group, we compute the median interval of last login time multiplied by the group size. If the value is much smaller than that of a normal group, users in the group login in group with the similar time, and the group is tracked as the Sybil group.

Tracking Using IP Address

We utilize IP addresses when users register their accounts. If suspicious users are registered with similar IP addresses, they are likely to be controlled by the same attacker. We classify suspicious users based on prefixes of their IP addresses. For example, the suspicious user has the IP address as A.B.C.D. We extract the prefix as A.B, and put the user into the corresponding group. All suspicious users are divided into groups by their IP addresses. Attackers often create many sybil users and control them to collude together. Small groups are useless for attacks.

Detecting Using their Activities

If a Sybil user posts a spammed content the application identifies him and coins him a Sybil after browsing through the user and his connections' identity. Likewise, even if a legitimate user posts a spam content by chance, the application doesn't mark him as Sybil as it checks the history and the status of the users' connections. With help of this we can easily track or detect the Sybil groups or users.

Validating Sybil Users/Groups

To validate Sybil groups, we use the widely acknowledged distinguishing feature: the action time similarity [10]. The action time similarity is based on the intuition that all users in a Sybil group take coordinated actions within the similar time. For example, many users post spam in a short time in Facebook, and their posting time is similar. In order to save the time cost, the attacker simultaneously controls all users in a Sybil group to take actions.

We utilize the median interval of action time to measure the similarity [17]. We summarize the computation of median interval in algorithm. Firstly, we sort action time of all users in a group. Then we measure the absolute time

interval between consecutive actions, and extract the median value of all such intervals. If users take actions within the short time, their action time is similar and the median interval is small; if users randomly take actions within the long time, their action time is dissimilar and the median interval is large [9].

The median interval characterizes the time similarity of actions taken by all users in a group. However, the median interval is negatively correlated with the group size. The bigger the group is, the more intensively user actions are distributed in the time period, which causes the smaller median interval. In contrast, the small group is likely to have the large median interval. In order to reduce the impact of the group size, we multiply the median interval by the group size.

We Define Three Validation Methods

The Validation Based on Last Login Time

The last login time describes the time when the user logins into the OSN for the last time. For a group, we compute the median interval of last login time multiplied by the group size. If the value is much smaller than that of a normal group, users in the group login in group with the similar time, and the group is validated as the Sybil group.

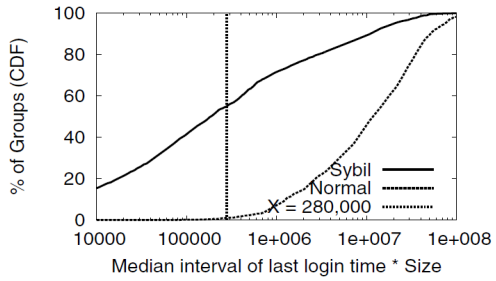

Figure 2: Group Distribution of Sybil Groups and Normal Groups

The Validation Based on Register Time

The register time is defined as the time when the user registers a new account in the system. For a group, we compute the median interval of register time multiplied by the group size. If the value is much smaller than that of a normal group, the register time of users are distributed too intensively, and we validate the group as the Sybil group.

The Validation Based on Friend Establishment Time

The friend establishment time describes the time when the user establishes the social relationship with a friend. For a group, we compute the median interval of friend establishment time multiplied by the number of friend relationships. Note that we mainly consider friend relationships within the group. The attacker simultaneously controls sybil users to make friends with others in the same group. Therefore, establishment time of social relationships within the group is likely to be similar. Sybil users also send friend requests to normal users outside of their group. The friend establishment time is determined by normal users, instead of attackers. If the result is extremely smaller than that of a normal group, social relationships within the group are established within the short time, and their group is validated as the sybil group.

CONCLUSIONS

I present the first attempt to identify and validate Sybil groups in the real system. First of all, we build the Sybil group detector based on multiple attributes, including their activities, popularity, social degree, friend relationship and IP

address. We design automatic validation mechanisms of Sybil users/groups. Our Sybil group detection and validation mechanisms have important implications for system design to defend against Sybil attacks in OSNs.

REFERENCES

1. Jing Jiang, Zifei Shan, Wenpeng Sha, Xiao Wang, Yafei Dai, "Detecting and Validating Sybil Groups in theWild," in 32 International conference on distributing computer system year 2012
2. J. Ratkiewicz, M. D. Conover, B. M. Meiss, Goncalves, A. Flammini, and F. Menczer, "Detecting and tracking political abuse in social media," in *Proc. of ICWSM*, July 2011.
3. H. Gao, J. Hu, C. Wilson, Z. Li, Y. Chen, and B. Y. Zhao, "Detecting and characterizing social spam campaigns," in *Proc. of ACM Internet Measurement Conference*, November 2010, pp. 35–47.
4. C. Grier, K. Thomas, V. Paxson, and M. Zhang, "@spam: The underground on 140 characters or less," in *Proc. Of ACM Conference on Computer and Communications Security*, October 2010, pp. 27–37.
5. D. Irani, SteveWebb, and C. Pu, "Study of static classification of social spam profiles in myspace," in *Proc. of ICWSM*, May 2010.
6. G. Stringhini, C. Kruegel, and G. Vigna, "Detecting spammers on social networks," in *Proc. of Annual Computer Security Applications Conference*, December 2010, pp. 1–9.
7. K. Thomas, C. Grier, V. Paxson, and D. Song, "Suspended accounts in retrospect: An analysis of twitter spam," in *Proc .of ACM Internet Measurement Conference*, November 2011.
8. M. Kirkpatrick, "Social networking now more popular than email, report finds," Read Write Web, March 2009.
9. H. Yu, M. Kaminsky, P. B. Gibbons, and A. D. Flaxman, "Sybilguard: Defending against sybil attacks via social networks," *IEEE/ACM Transactions on Networking*, vol. 16, no. 3, pp. 576–589, June 2008.
10. H. Yu, P. B. Gibbons, M. Kaminsky, and F. Xiao, "Sybillimit: A near-optimal social network defense against sybil attacks," in *Proc. of the IEEE Symposium on Security and Privacy*, May 2008, pp. 3–17.
11. G. Danezis and P. Mittal, "Sybilinfer: Detecting sybil nodes using social networks," in *Proc. of NDSS*, February 2009.
12. N. Tran, B. Min, J. Li, and L. Subramanian, "Sybil-resilient online content voting," in *Proc. of NSDI*, April 2009, pp. 15–28.
13. Z. Yang, C. Wilson, X. Wang, T. Gao, B. Y. Zhao, and Y. Dai, "Uncovering social network sybils in the wild," in *Proc. Of ACM Internet Measurement Conference*, November 2011.
14. J. Jiang, C. Wilson, X. Wang, P. Huang, W. Sha, Y. Dai, and B. Y. Zhao, "Understanding latent interactions in online social networks," in *Proc. of ACM Internet Measurement Conference*, November 2010, pp. 369–382.
15. B. Viswanath, A. Post, K. P. Gummadi, and A. Mislove, "An analysis of social network-based sybil defenses," in *Proc. Of SIGCOMM*, August 2010, pp. 363–374.

16. A. Mohaisen, N. Hopper, and Y. Kim, "Keep your friends close: Incorporating trust into social network-based Sybil defenses," in *Proc. of the IEEE Infocom*, April 2011.

17. K. Lee, J. Caverlee, and S. Webb, "Uncovering social spammers: Social honeypots + machine learning," in *Proc. Of SIGIR*, July 2010, pp. 435–442.

18. Z. Chu, S. Gianvecchio, H. Wang, and S. Jajodia, "Who is tweeting on twitter: Human, bot, or cyborg?" in *Proc. of Annual Computer Security Applications Conference*, December 2010, pp. 21–30.

19. S. Webb, J. Caverlee, and C. Pu, "Social honeypots: Makingfriends with a spammer near you," in *Proc. of CEAS*, Mountain View, USA, August 2008.

20. F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida, "Detecting spammers on twitter," in *Proc. of CEAS*, Washington, USA, July 2010.

21. F. Benevenuto, T. Rodrigues, V. Almeida, J. Almeida, and M. Gonglves, "Detecting spammers and content promoters in online video social networks," in *Proc. of SIGIR*, Boston, USA, July 2009.

22. S. Yardi, D. Romero, G. Schoenebeck, and D. boyd, "Detecting spam in a twitter network," *First Monday*, vol. 15, no. 1-4, January 2010.